doi:10.1093/scan/nsr008 SCAN (2012) 7, 282–288

Disrupting the right prefrontal cortex alters moral judgement

Sébastien Tassy,[1,2] Olivier Oullier,[3] Yann Duclos,[4] Olivier Coulon,[5] Julien Mancini,[2,6] Christine Deruelle,[1] Sharam Attarian,[2,7] Olivier Felician,[2,8] and Bruno Wicker[1,9]

[1]Institut de Neurosciences Cognitives de la Méditerranée, CNRS UMR 6193 & Université de la Méditerranée, Marseille, France, [2]Assistance Publique-Hôpitaux de Marseille, [3]Laboratoire de Psychologie Cognitive, UMR 6146, Université de Provence & CNRS, Aix-Marseille Université, Marseille, France, [4]Plasticité et Physio-pathologie de la Motricité, UMR 6196 CNRS & Université de la Méditerranée, Marseille, [5]Laboratoire des Sciences de l'Information et des Systèmes, UMR 6168 CNRS, Marseille, [6]Laboratoire d'Enseignement et de Recherche sur le Traitement de l'Information Médicale, Faculté de Médecine, Université de la Méditerranée, Marseille, [7]Laboratoire de Génétique Médicale et Fénomique Fonctionnelle, INSERM U910, Faculté de Médecine, Université de la Méditerranée, Marseille, [8]Laboratoire Epilepsie et Cognition, INSERM U 751, Faculté de Médecine, Université de la Méditerranée, Marseille, France, and [9]Integrative Neurosciences Laboratory, Physics Department, University of Buenos Aires, Pabellón I, Ciudad Universitaria, (1428) Capital Federal, Argentina

Humans daily face social situations involving conflicts between competing moral decision. Despite a substantial amount of studies published over the past 10 years, the respective role of emotions and reason, their possible interaction, and their behavioural expression during moral evaluation remains an unresolved issue. A dualistic approach to moral evaluation proposes that the right dorsolateral prefrontal cortex (rDLPFc) controls emotional impulses. However, recent findings raise the possibility that the right DLPFc processes emotional information during moral decision making. We used repetitive transcranial magnetic stimulation (rTMS) to transiently disrupt rDLPFc activity before measuring decision making in the context of moral dilemmas. Results reveal an increase of the probability of utilitarian responses during objective evaluation of moral dilemmas in the rTMS group (compared to a SHAM one). This suggests that the right DLPFc function not only participates to a rational cognitive control process, but also integrates emotions generated by contextual information appraisal, which are decisive for response selection in moral judgements.

Keywords: moral judgement; rTMS; right prefrontal cortex; emotion; utilitarism; decision

INTRODUCTION

The profound nature of moral judgement has been discussed for centuries across several fields including philosophy and more recently psychology. The Kantian tradition supported by developmental psychologists such as Kohlberg (1969), conceives moral judgement as the product of conscious, effortful rational reasoning. A different view is expressed in Hume's so-called 'sentimentalism', which emphasizes the role of more intuitive and affective reactions as a guide to moral judgement (Haidt, 2007). By considering both reason and emotion as important forces driving moral judgement, recent research in neuroscience has led to reconcile these opposite traditions and to propose a model relying on a dualistic cerebral *modus operandi* (Green *et al.*, 2001). Among the major findings is the fact that the right dorsolateral prefrontal cortex (rDLPFc) could play an important role in cognitive 'rational' control processes leading to utilitarian judgements. The medial prefrontal structures, on the other hand, would be responsible for more intuitive emotional reactions (Green *et al.*, 2004; Koenigs *et al.*, 2007). However, in the field of social cognition, a role of the right DLPFc restricted to rational cognitive control has recently been questioned in a study requiring healthy participants to make decisions in the ultimatum game (UG), a form of economic social dilemma generating a conflict between reason (accepting unfair monetary offers from someone else) and emotion (rejecting unfair offers) (Sanfey *et al.*, *2003*; Knoch *et al.*, 2006). Low-frequency repetitive transcranial magnetic stimulation (rTMS) pulses delivered over the right DLPFc substantially reduced subjects' willingness to reject their partners' intentionally unfair offers, suggesting a reduction of an anger impulse, while still being able to judge such offers as unfair. In other words, disruption of the right DLPFc, a so-called 'rational' part of the brain, favoured what economists often consider a rational behaviour. This finding refutes a functional role of the right DLPFc solely restricted to 'rational' cognitive control over intuitive affective reactions (Moll and de Oliveira-Souza, 2007; Talmi and Frith, 2007). Among the proposed alternate hypotheses, Moll suggested that emotions in play during moral evaluation should be divided in different subcomponents that

Received 13 August 2010; Accepted 1 February 2011
Advance Access publication 22 April 2011
We thank the medical imaging department of Marseille Laveran Military Hospital for T1 MRI acquisitions.
Correspondence should be addressed to Bruno Wicker. Institut de Neurosciences Cognitives de la Méditerranée, CNRS UMR 6193 and Université de la Méditerranée, 31 chemin Joseph Aiguier, 13402 Marseille cedex 20, France. E-mail: wicker@incm.cnrs-mrs.fr

would be processed by two different cortical systems. In line with this idea, we propose that disruption of the right DLPFc may have specifically prevented the integration of complex emotional responses generated by appraisal of contextual information such as sophisticated fairness norm, and led to a reduction of the emotional weight during decision making in social contexts.

The present study investigates whether this hypothesis stands in the context of moral dilemmas designed to pit two competing moral considerations against one another: an utilitarian consideration maximizing the aggregate welfare by acting at one's (or a few individuals') expense, and a more 'emotional' consideration reflecting a strong emotional intuitive and automatic aversion to the proposed action. If, as we hypothesize, the right DLPFc plays a role in the integration of emotional responses generated by appraisal of complex social information we expect an increased probability of utilitarian responses. This reduction would be the consequence of a reduced emotional bias during response selection. In contrast, a dualistic cerebral process hypothesis would predict a reduction of the rational control over emotional impulses, hence a reduced probability of utilitarian responses.

Inspired by findings of divergent effects of disrupting the activity of the right DLPFc on fairness judgement and effective choice in the UG, and since psychopathic patients exhibit inappropriate moral behaviour in spite of intact moral judgement (Cima *et al.*, 2010), our experimental design dissociates two aspects of moral evaluation. One is a judgement of acceptability (i.e. 'Is this action morally permissible?') that involves an impersonal objective evaluation. The other is a first-person perspective subjective 'choice' (*Would you perform this action?*). Given the role of the lateral prefrontal cortex in objective contextual cues integration, our secondary hypothesis is that disruption of its activity may influence differently the objective judgement task and the subjective behavioural choice (Kouneiher *et al.*, 2009).

METHODS

Participants

Twenty-four right-handed men, without any antecedent of either neurological or psychiatric disease, volunteered in the study. Each of them participated to only one of the two experimental conditions (rTMS over the right DLPFc, $n = 12$, or SHAM stimulation over the same location, $n = 12$), and none had previously experienced TMS, rTMS or moral dilemma experiments. There was no significant difference between groups with respect to age, education, number of siblings, mother tongue and religion. All participants signed informed consent and underwent a medical safety screening by a physician according to international safety guidelines for the use of TMS. They received a monetary compensation for participating to the experiment. The experiment received full approval of the local IRB (Comité de Protection de la Personne Sud-Méditerranée II).

TMS

Neuronatomical localization

A T1-weighted MRI was acquired for all participants to exactly localize the optimal stimulation sites for rTMS. The right DLPFc was identified based on the Talairach coordinates $x = 45$, $y = 36$, $z = 24$. These coordinates are those of the right DLPFc region involved in moral dilemma resolution as reported by Greene *et al.* (2001). For correct placement of the TMS coil in space, we used the eXimia Navigated Brain Stimulation system 2.3 (Nexstim, Helsinki, Finland). We transformed the coordinates to each participant's native brain space using AFNI software (http://afni.nimh.nih.gov/afni/), in order to make them suitable for neuronavigation. For each of them, the voxel corresponding to the coordinates of the right DLPFc was identified. A virtual functional map was then created by building an image of the same size and resolution than the T1 image with a null value everywhere, and to which a 10-mm Gaussian kernel centred on the right DLPFc voxel was added. This virtual functional map was then used as a target by the neuronavigation system.

Stimulation

rTMS was delivered using a MagPro X100 system with option (MagVenture, Farum, Denmark) and a MCF-B65 butterfly coil (75-mm diameter double circle, air-cooled) for stimulation. Stimulation parameters were identical to those used by Knoch and colleagues in the dilemma game experiment cited above and are supposed to inhibit the targeted area (Knoch *et al.*, 2006). Stimulation intensity was therefore set to 54% of the stimulator maximum output. The coil was held tangential to the subject's head with the handle pointing rostrally. Participants received a single, 15-min, 1-Hz rTMS train (900 pulses) over the right DLPFc (rTMS group, $n = 12$), or a sham stimulation with a MCF-P-B65 placebo butterfly coil (sham group, $n = 12$) of the same duration. The rTMS parameters were well within currently recommended guidelines (Rossi *et al.*, 2009), resulting in an estimated modification of DLPFc perfusion for 14 min (Eisenegger *et al.*, 2008) and in a reduction of excitability of the targeted cortical region for several minutes following completion of the rTMS train (Maeda *et al.*, 2000).

Experimental design

Stimuli

The dilemma scenarios were selected from a battery developed by Greene and colleagues (2001), from the *Moral sense test* of Marc Hauser's lab at Harvard University, or created by the experimenters. Dilemmas in English were translated to French and adapted to take into account cultural specificities. In a moral dilemma, the participant faces a conflict between two opposing moral values or requirements, while a non-moral dilemma opposes situations with no moral connotation.

For the purpose of post hoc analysis, an ancillary behavioural experiment explored the emotional value of each

dilemma by having an independent group of 52 male students rating the level of emotional intensity of each selected scenario. This allowed studying a potential correlation between the emotional level of each dilemma and the probability of utilitarian response. We also identified the level of conflict generated by each dilemma, using the methodology previously employed by Greene *et al.* (2004). Three groups of five dilemmas were defined according to response time in the SHAM group: low-conflict = 5 dilemmas with lowest mean response time, mild-conflict = 5 intermediate mean response time dilemmas, and high-conflict = 5 dilemmas with longest mean response time.

Experimental protocol

Participants received instructions explaining the task prior to stimulation. Right after real or sham stimulations, participants faced the scenarios of 15 moral and 9 non-moral dilemmas. After reading a scenario, each participant was required to answer a first question 'would it be *acceptable* to . . . ?' (Objective judgement task) by yes or no. Then the same scenario was presented again and the participant was required to answer a second question 'would *you do* . . . ?' (subjective choice task) by yes or no. Moral and non-moral dilemmas were presented on a computer screen in a randomized order using the Presentation software. Subjects gave their answer by pressing one of two mouse buttons corresponding to the two answers (yes or no), placed randomly on the right or left side of the computer screen. Participants were instructed to imagine each hypothetical situation as vividly and realistically as they could, and could read and respond at their own pace.

Responding to all dilemmas took an average duration of 11.3 ± 2.6 min, which is shorter than the effect of rTMS stimulation on rDLPFc perfusion demonstrated by Eisenegger and colleagues using the same parameters (single, 15-min, 1-Hz rTMS train, 900 pulses) (Eisenegger *et al.*, 2008).

Statistical analyses

Each response was coded 1 when it favoured maximizing the good of more people at the expense of very few identified individuals ('utilitarian' response maximizing aggregate welfare e.g. sacrificing one person's life to save five), and 0 for the reverse situation. For non-moral dilemmas, 'appropriate' responses implied the maximization of beneficence overall consequences (e.g. buying a new television instead of repairing the old one for the same price was coded 1 and 'inappropriate' was coded 0).

All statistical analyses were carried out using the Statistical Package for the Social Sciences (SPSS, SPSS Inc., Chicago). To test for between-group differences in the probability of utilitarian responses in both tasks we used a logistic regression fitted with the generalized estimating equations method (Keonigs *et al.*, 2007). Correlation between the probability of utilitarian responses and the level of emotion assigned to a dilemma was tested using Pearson correlation coefficient.

RESULTS

Non-moral dilemmas

Responses are identical in the rTMS and SHAM groups for non-moral dilemmas (judgement $P > 0.38$; choice $P > 0.65$), showing that disruption of right DLPFC activity did not affect the ability to evaluate non-moral conflicts and make appropriate behavioural choices.

Moral dilemmas

Objective judgement task

Probability of utilitarian responses in the judgement task was significantly higher in the rTMS group [rTMS $= 0.69 \pm 0.03$; SHAM $= 0.57 \pm 0.04$; odds ratio (OR) $= 1.70$; $P = 0.031$; Figure 1]. For instance, during the classic trolley dilemma, subjects of the rTMS group would typically more often judge acceptable to deviate the trolley in order to save five people at the expense of one. Disruption of right DLPFc activity thus led to a significantly higher level of utilitarianism. An increase of utilitarian response is classically associated to a reduced influence of emotion. To explore the possibility that this effect of rTMS disruption reflects a reduced influence of emotion during evaluation of a dilemma, we performed a post hoc analysis studying the correlation between the level of emotional intensity of a dilemma (cf. Methods section) and the probability of utilitarian response. In the SHAM group, there is a negative correlation between the probability of utilitarian response and the level of emotional intensity of a dilemma (Pearson correlation coefficient, $r = -0.565$; $P = 0.028$) (cf. Figure 2). No correlation between the probability of utilitarian response and the level of emotional intensity of a dilemma is observed in the rTMS group ($r = 0.292$; $P = 0.29$) (cf. Figure 2).

This absence of correlation between probability of utilitarian response and the level of emotional intensity of a dilemma in the rTMS group suggests that disruption of right DLPFc activity led to a disturbed appraisal of the emotional content of dilemmas. The right DLPFc may thus be involved in the processing of the emotional value of a dilemma.

Subjective choice task

Analyses of responses provided in the choice task revealed no significant statistical difference between groups regarding the probability of utilitarian response (rTMS $= 0.62 \pm 0.04$; SHAM $= 0.72 \pm 0.03$, OR $= 0.625$, $P > 0.059$; and non-moral $P > 0.65$) (cf. Figure 3). However, a trend towards reduced probability of utilitarian responses was observed in the rTMS group ($P = 0.06$). The effect is of the same magnitude, but opposite to the effect observed for the judgement task.

The role of right DLPFc activity on biasing decision towards more utilitarianism has been previously described, but only in high-conflict dilemmas evaluation requiring first-person perspective decision that parallels our choice

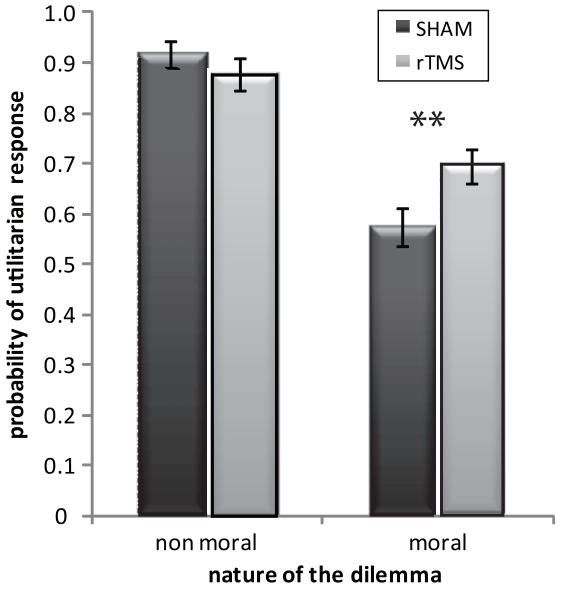

Fig. 1 Probability of utilitarian response in the Objective judgement task. Disruption of the right DLPFc by low-frequency rTMS led to a higher probability of utilitarian responses (between-group comparison: SHAM/rTMS, logistic regression fitted with the GEE, OR = 1.70; $^{**}P = 0.031$).

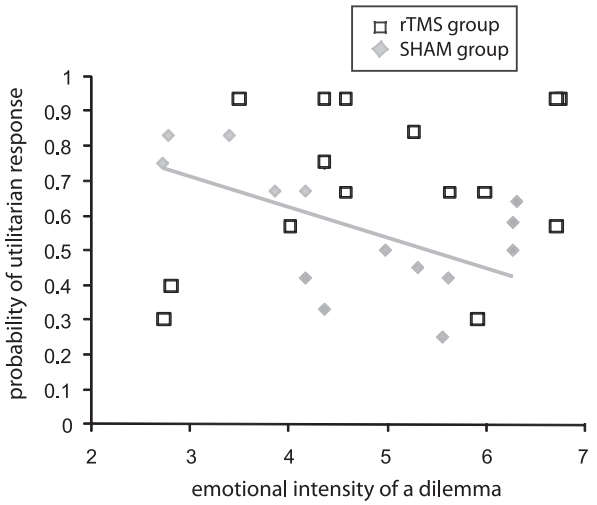

Fig. 2 Emotional bias in the Objective judgement task the probability of utilitarian responses decreases when the emotional intensity attributed to the dilemma increases in the SHAM group (Pearson correlation coefficient, $r = -0.565$; $P = 0.028$), but not in the rTMS group (Pearson correlation coefficient $r = 0.292$; $P = 0.29$).

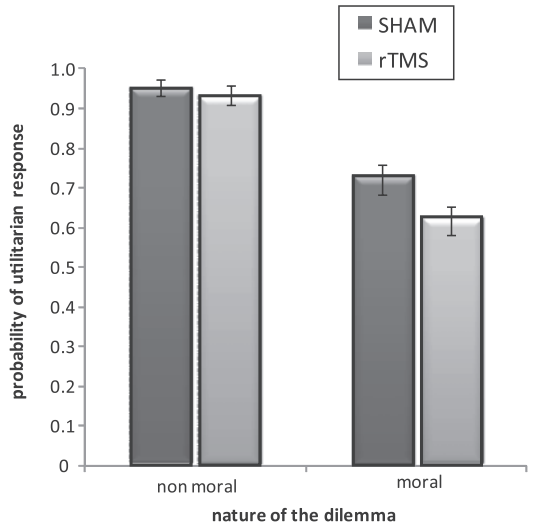

Fig. 3 Probability of utilitarian response in the subjective choice task (between-group comparison: SHAM/rTMS, logistic regression fitted with the GEE, OR = 0.625, $P > 0.059$).

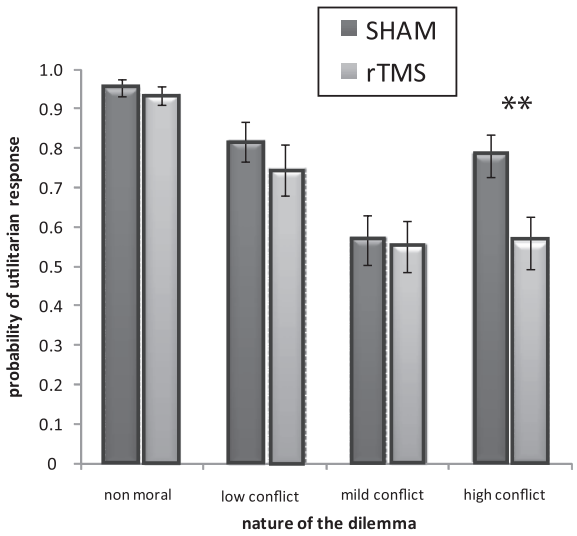

Fig. 4 Probability of utilitarian response in the subjective choice task, depending on the level of internal conflict generated by a dilemma. The anti-utilitarian effect of rTMS is present in the case of high-conflict dilemmas only (between group comparison: SHAM/rTMS, logistic regression fitted with the GEE, OR = 0.248; $^{**}P = 0.032$).

task (Greene *et al.*, 2004). We thus conducted a post hoc analysis studying the responses in each category of dilemma ('low-conflict', 'mild-conflict' and 'high-conflict'). rTMS over the right DLPFc decreased significantly the probability of utilitarian response in the choice task for high-conflict (OR = 0.248; $P = 0.032$), but not for 'low-conflict' and 'mild-conflict' dilemmas (cf. Figure 4). Excluding high-conflict dilemmas from the analysis totally suppressed the trend towards less utilitarianism in the rTMS group (OR = 0.813; $P = 0.467$). The anti-utilitarian effect of rTMS on right DLPFc activity thus seems specific to high-conflict dilemmas.

DISCUSSION

Our study explored the consequence of a localized disruption of right DLPFc activity in the context of moral dilemma evaluation. Moral evaluation was tested in two different conditions, with either a judgement of general objective acceptability or a first person subjective endorsement. None of these conditions can be considered as implying a decision making process with direct effective consequences for the

subject, such as in economic decision making for example. However, both conditions involved a mechanism for computing the value of two options and identifying the most acceptable of these options to reach a complex social decision.

Objective judgement task

Disruption of rDLPFc activity affected subjects' behaviour by increasing the probability of utilitarian responses. This confirms a role of the right DLPFc in moral judgement, in line with a growing body of data on decision making in complex social contexts reporting that lateral prefrontal structures play a role in valuation processes by integrating various kinds of inputs (Kable and Glimcher, 2009; Forbes and Grafman, 2010; Rangel and Hare, 2010). More specifically, our results reveal that the right DLPFc has a role in computing the value of attributes such as the impact of a choice on others' well being.

Disrupting rDLPFc activity led to a higher level of utilitarianism. This increased probability of utilitarian response does not fit with the dual system hypothesis predicting that right DLPFc should code for 'rational' cognitive control over emotional impulse. On the contrary, similarly to Knoch *et al.* 2006 results in the context of UG, it shows that right DLPFc activity is not responsible for the utilitarian bias. This has previously led some authors to propose that the right DLPFc may be part of a psychological system that participates to the integration of representational emotions during moral evaluation (Moll *et al.*, 2005; Cushman *et al.*, 2011). Result of our post hoc analyses suggesting a causal link between rTMS effect and the emotional value of a dilemma support this view. The question remains of what types of process subserved by right DLPFc activity are directly related to the reduction of emotional impulse during objective evaluation of moral dilemmas. Other authors have suggested that the right DLPFc would be important for self-centred and other-aversive emotional experience in social context (Moll and de Oliveira-Souza, 2007) or would be involved in strategic planning of the emotional response through currency emotions designed to participate in the process of practical reasoning, providing negotiable motivations for and against different behaviours (Cushman *et al.*, in press). In our view, the right DLPFc would code for secondary emotions necessary to implement behaviour relying on external guidance (Moll *et al.*, 2005) and abstract rules processing (Passingham *et al.*, 2010). This role of the right DLPFc in objective integration of external cues contrasts with a role of medial prefrontal cortex in processing the subjective values of actions (Kouneiher *et al.*, 2009; Seitz *et al.*, 2009; Passingham *et al.*, 2010). It would explain right DLPFc activation during judgements involving reference to sophisticated social norms (Buckholtz *et al.*, 2008) and its involvement in processing fairness norms (Knoch *et al.*, 2006). This function of right DLPFc could explain why individuals suffering ventromedial prefrontal cortex (VMPFc) lesions to make 'over-emotional' decisions (economically speaking) when a sophisticated abstract norm of fairness is breached in the UG (Koenigs and Tranel, 2007).

Subjective choice task

Although not statistically significant, the effect of rTMS over the right DLPFc evoked a trend towards less utilitarianism. This effect is of comparable magnitude but opposite to the one observed in the objective judgement task. Our secondary hypothesis favouring a different effect for rDLPFc disruption in the objective and subjective evaluation tasks is therefore validated: evaluative judgement and first-person subjective behavioural choice may rely on relatively independent processes at the neurobiological level.

Interestingly, post hoc analysis revealed that the rTMS effect in the subjective choice task is specific to high-conflict dilemmas (cf. Figure 4). This is in line with a previous neuroimaging study reporting that right DLPFc activation correlates with a bias in the response towards more utilitarianism only in the case of high-conflict moral dilemmas (Greene *et al.*, 2004). A tentative explanation is that the right DLPFc may be a part of the neural system responsible for rational control over prepotent affective response, but only during subjective evaluation of highly conflicting moral dilemmas. In this case, the conflict may not be generated by competition between reason and emotion but rather by an interaction between two emotional subprocesses, as proposed by Moll and de Oliveira-Souza (2007). The bias towards utilitarianism found in high-conflict dilemmas in the subjective evaluation task would be the consequence of the integration by the right DLPFc of secondary social emotions competing with more automatic/intuitive ones, rather than result from the overriding of intuitive emotional reactions by reason. This is illustrated in the well known high-conflict 'crying baby' dilemma, where the choice is given to a mother to either accept to kill her own child to save many people, or to refuse and be responsible for the death of all of these people. Here, the influence of the very intuitive visceral emotional reaction that stems from imagining killing one's own baby would compete with the secondary social emotions induced by planning future consequences of inaction, (namely to the death of all), and pure rational processes restricted to arithmetic computation of saving no *vs* many lives (Shenhav and Greene, 2010). In this case of high-conflict dilemma, the more activated the right DLPFc, the more secondary social emotions resulting from general objective contextual evaluation would compete with intuitive emotional reactions, biasing responses towards a more utilitarian decision. In less conflicting dilemmas, the two kinds of emotions would bias responses similarly towards less utilitarianism. Thus, disrupting the automatic/intuitive emotional system would result in an increase of utilitarian responses only in high-conflict dilemmas. This interpretation is in line with Koenig's *et al.*' findings when exploring moral evaluation of patients with damage to the

prefrontal cortex (Koenigs *et al.*, 2007). It also explains that in non-conflicting dilemmas, the two types of emotions would have influenced along the same direction, biasing responses towards less utilitarianism.

Since refusing to kill one's own baby can be considered as a form of selfishness, this result also appears in line with findings from a previous study hypothesizing that right DLPFc disruption frees selfish impulses when facing unfairness in the UG (Knoch *et al.*, 2006). Complex secondary emotions would compete with more automatic 'selfish' processes (either purely rational or emotional), generating a strong internal conflict. Cognitive control would thus reduce influence of intuitive emotional reaction not just by 'attenuation of prepotent emotion intensity' but by generation of competing secondary emotions or motivations.

Noteworthy, the effect of rTMS in the subjective task is opposite to the one observed on responses in the objective judgement task. This opposite result in the two tasks raise the possibility that Moll's proposition of emotions being divided into two subcomponents as well as Greene's proposition of a dual model are complementary, and support the hypothesis that three distinct processes would interplay during decision making in moral dilemma: rational computation, intuitive emotional reactions and secondary emotions. Purely rational processes restricted to arithmetic computation of saving no *vs* many lives would interplay either with automatic/intuitive emotions and secondary social emotions. Imbalance between these processes and their respective coordinated influence on moral decision making would depend on the context in which the dilemma occurs, such as when one engages in evaluative judgement or subjective behavioural choice. In the case of objective evaluation, moral decision would rely on the interplay between pure rational processes and mainly secondary social emotions. The influence of the latter (coded by the rDLPFC) is crucial in this case and their suppression would have an overall utilitarian effect, as observed in our results. Such evaluative judgements are probably the ones that specifically activated right DLPFc during impersonal moral dilemma as defined by Greene *et al.* (2001). On the other hand, in the case of subjective evaluation, moral decision would rely on the interplay between purely rational processes and mainly automatic/intuitive emotions. The influence of secondary social emotions (coded by the rDLPFc) would be marginal here and their suppression have an overall marginal effect. A case not to oversee would however be when secondary social emotions strongly compete with automatic emotional impulse, leading to so called 'high-conflict' dilemmas. In these dilemmas, the influence of secondary social emotions deeply modifies the equilibrium between rational processes and emotional impulse. Their suppression slightly increases influence of the automatic/basic emotional impulse towards less utilitarianism, as in the results of the present study.

Overall, our results suggest that the rDLPFc, most likely in combination with other brain structures, integrates complex secondary emotions generated by contextual information knowledge that are decisive for response selection in moral judgement. This nicely illustrates Spinoza's early intuition about the necessary role of '*Emotions which are aroused or spring from reason*' for an Ethic to exist. Morality would thus result from the interplay (collaborative or competitive) between automatic emotional reaction, emotional contextual appraisal and rational welfare maximization. This finding is in line with a more general evolutionist view of the psychology of decision making in the context of social dilemmas where three neural mechanisms could coexist: a more primitive one that would favour direct self-interest; a more automatic favouring welfare of kin/in-group members; and a third more conscious mechanism implementing norms beneficial to large social groups. The influence of frames of reference (objective *vs* subjective) and context (low *vs* highly conflicting situations) on a subtle equilibrium between these three processes is probably one of the key features to elucidate any 'moral grammar' (Hauser, 2007).

Conflict of Interest

None declared.

REFERENCES

Buckholtz, J.W., Asplund, C.L., Dux, P.E., et al. (2008). The neural correlates of third-party punishment. *Neuron*, *60*, 930–40.

Buckley, M.J., Mansouri, F.A., Hoda, H., et al. (2009). Dissociable components of rule-guided behavior depend on distinct medial and prefrontal regions. *Science*, *325*, 52–8.

Cima, M., Tonnaer, F., Hauser, M.D. (2010). Psychopaths know right from wrong but don't care. *Social Cognitive and Affective Neuroscience*, *5*, 59–67.

Cushman, F., Young, L., Greene, J.D. (2011). Our mutli-system moral psychology: towards a consensus view. In: Doris, J., Harman, G., Nichols, S., Prinz, J., Sinnott-Armstrong, W., Stich, S., editors. *The Oxford Handbook of Moral Psychology*. Oxford: Oxford University Press.

Eisenegger, C., Treyer, V., Fehr, E., Knoch, D. (2008). Time-course of 'off-line' prefrontal rTMS effects–a PET study. *Neuroimage*, *42*, 379–84.

Forbes, C.E., Grafman, J. (2010). The role of the human prefrontal cortex in social cognition and moral judgment. *Annual Review of Neuroscience*, *33*, 299–324.

Greene, J.D., Nystrom, L.E., Engell, A.D., Darley, J.M., Cohen, J.D. (2004). The neural bases of cognitive conflict and control in moral judgment. *Neuron*, *44*, 389–400.

Greene, J.D., Sommerville, R.B., Nystrom, L.E., Darley, J.M., Cohen, J.D. (2001). An fMRI investigation of emotional engagement in moral judgment. *Science*, *293*, 2105–8.

Haidt, J. (2007). The new synthesis in moral psychology. *Science*, *316*, 998–1002.

Hauser, M. (2007). *Moral Minds: How Nature Designed Our Universal Sense of Right and Wrong*. New York: HarperCollins.

Kable, J.W., Glimcher, P.W. (2009). The neurobiology of decision: consensus and controversy. *Neuron*, *63*, 733–45.

Koenigs, M., Tranel, D. (2007). Irrational economic decision-making after ventromedial prefrontal damage: evidence from the Ultimatum Game. *Journal of Neuroscience*, *27*, 951–6.

Koenigs, M., Young, L., Adolphs, R., et al. (2007). Damage to the prefrontal cortex increases utilitarian moral judgements. *Nature*, *446*, 908–11.

Knoch, D., Pascual-Leone, A., Meyer, K., Treyer, V., Fehr, E. (2006). Diminishing reciprocal fairness by disrupting the right prefrontal cortex. *Science*, *314*, 829–32.

Kohlberg, L. (1969). Stage and sequance: the cognitive-developmental approach to socialization. In: Goslin, D., editor. *Handbook of Socialization Theory and Research*. Chicago: Rand McNally, pp. 347–480.

Kouneiher, F., Charron, S., Koechlin, E. (2009). Motivation and cognitive control in the human prefrontal cortex. *Nature Neuroscience*, *12*, 939–45.

Maeda, F., Keenan, J.P., Tormos, J.M., Topka, H., Pascual-Leone, A. (2000). Modulation of corticospinal excitability by repetitive transcranial magnetic stimulation. *Clinical Neurophysioloy*, *111*, 800–5.

Moll, J., de Oliveira-Souza, R. (2007). Moral judgments, emotions and the utilitarian brain. *Trends in Cognitive Science*, *11*, 319–21.

Moll, J., Zahn, R., de Oliveira-Souza, R., Krueger, F., Grafman, J. (2005). Opinion: the neural basis of human moral cognition. *Nature Review. Neuroscience*, *6*, 799–809.

Passingham, R.E., Bengtsson, S.L., Lau, H.C. (2010). Medial frontal cortex: from self-generated action to reflection on one's own performance. *Trends in Cognitive Sciences*, *14*, 16–21.

Rangel, A., Hare, T. (2010). Neural computations associated with goal-directed choice. *Current Opinion on Neurobiology*, *20*, 262–70.

Rossi, S., Hallett, M., Rossini, P.M., Pascual-Leone, A. (2009). Safety, ethical considerations, and application guidelines for the use of transcranial magnetic stimulation in clinical practice and research. *Clinical Neurophysiology*, *120*, 2008–39.

Sanfey, A.G., Rilling, J.K., Aronson, J.A., Nystrom, L.E., Cohen, J.D. (2003). The neural basis of economic decision-making in the ultimatum game. *Science*, *300*, 1755–8.

Seitz, R.J., Franz, M., Azari, N.P. (2009). Value judgments and self-control of action: the role of the medial frontal cortex. *Brain Research Review*, *60*, 368–78.

Shenhav, A., Greene, J.D. (2010). Moral judgments recruit domain-general valuation mechanisms to integrate representations of probability and magnitude. *Neuron*, *67*, 667–77.

Talmi, D., Frith, C. (2007). Neurobiology: feeling right about doing right. *Nature*, *446*, 865–6.